

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Petkovich, P. Martin, White, Jay A.,  
Beckett, Barbara R., Jones, Glenville
- (ii) TITLE OF INVENTION: Retinoid Metabolizing Protein
- (iii) NUMBER OF SEQUENCES: 35
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Blake, Cassels & Graydon
  - (B) STREET: Box 25, Commerce Court West
  - (C) CITY: Toronto
  - (D) PROVINCE: Ontario
  - (E) COUNTRY: Canada
  - (F) ZIP: M5L 1A9
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
  - (B) COMPUTER: COMPAQ, IBM PC compatible
  - (C) OPERATING SYSTEM: MS-DOS 5.1
  - (D) SOFTWARE: WORD PERFECT
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBERS: 08/667,546; 08/724,466
  - (B) FILING DATE: June 21, 1996; October 1, 1996
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Hunt, John C.
  - (B) REGISTRATION NUMBER: 36,424
  - (C) REFERENCE/DOCKET NUMBER: 50767/00010
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (416) 863-4344
  - (B) TELEFAX: (416) 863-2653

(2) INFORMATION FOR SEQ ID NO:1

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 337 base pairs
  - (B) TYPE: nucleic acid
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1

TGCCAGTGGG	CAATCTCCCT	ACCAAATTCA	CTAGTTATGT	CCAGAAATTA	GCCTAAACCG	60
GAGCCTTTGT	ACATATGTTT	TTATTTTAGA	TGAACTGTGA	TGTATTGGAT	ATTTTCTAAT	120
TTGTTTATAT	AAAGCAGATG	TGTATATAAG	TCTATGCGAA	GAAGCGAAAA	CGAGGGCACT	180
ACTTTCTCAT	GGATCACTGT	AATGCTACAG	AGTGTCTGTG	ATGTATATTT	ATAATGTAGT	240
TGTGTCATAT	AGCTTTTGTA	CTGTATGCAA	CTTATTTAAC	TCGCTCTTTA	TCTCATGGGT	300
TTTATTTAAT	AAAACATGTT	CTTACAAAAA	AAAAAAA			337

(2) INFORMATION FOR SEQ ID NO:2



(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 492 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2

Met	Gly	Leu	Tyr	Thr	Leu	Met	Val	Thr	Phe	Leu	Cys	Thr	Ile	Val	Leu	1	5	10	15
Pro	Val	Leu	Leu	Phe	Leu	Ala	Ala	Val	Lys	Leu	Trp	Glu	Met	Leu	Met	20	25	30	
Ile	Arg	Arg	Val	Asp	Pro	Asn	Cys	Arg	Ser	Pro	Leu	Pro	Pro	Gly	Thr	35	40	45	
Met	Gly	Leu	Pro	Phe	Ile	Gly	Glu	Thr	Leu	Gln	Leu	Ile	Leu	Gln	Arg	50	55	60	
Arg	Lys	Phe	Leu	Arg	Met	Lys	Arg	Gln	Lys	Tyr	Gly	Cys	Ile	Tyr	Lys	65	70	75	
Thr	His	Leu	Phe	Gly	Asn	Pro	Thr	Val	Arg	Val	Met	Gly	Ala	Asp	Asn	85	90	95	
Val	Arg	Gln	Ile	Leu	Leu	Gly	Glu	His	Lys	Leu	Val	Ser	Val	Gln	Trp	100	105	110	
Pro	Ala	Ser	Val	Arg	Thr	Ile	Leu	Gly	Ser	Asp	Thr	Leu	Ser	Asn	Val	115	120	125	
His	Gly	Val	Gln	His	Lys	Asn	Lys	Lys	Lys	Ala	Ile	Met	Arg	Ala	Phe	130	135	140	
Ser	Arg	Asp	Ala	Leu	Glu	His	Tyr	Ile	Pro	Val	Ile	Gln	Gln	Glu	Val	145	150	155	
Lys	Ser	Ala	Ile	Gln	Glu	Trp	Leu	Gln	Lys	Asp	Ser	Cys	Val	Leu	Val	165	170	175	
Tyr	Pro	Glu	Met	Lys	Lys	Leu	Met	Phe	Arg	Ile	Ala	Met	Arg	Ile	Leu	180	185	190	
Leu	Gly	Phe	Glu	Pro	Glu	Gln	Ile	Lys	Thr	Asp	Glu	Gln	Glu	Leu	Val	195	200	205	
Glu	Ala	Phe	Glu	Glu	Met	Ile	Lys	Asn	Leu	Phe	Ser	Leu	Pro	Ile	Asp	210	215	220	
Val	Pro	Phe	Ser	Gly	Leu	Tyr	Arg	Gly	Leu	Arg	Ala	Arg	Asn	Phe	Ile	225	230	235	
His	Ser	Lys	Ile	Glu	Glu	Asn	Ile	Arg	Lys	Lys	Ile	Gln	Asp	Asp	Asp	245	250	255	
Asn	Glu	Asn	Glu	Gln	Lys	Tyr	Lys	Asp	Ala	Leu	Gln	Leu	Leu	Ile	Glu	260	265	270	
Asn	Ser	Arg	Arg	Ser	Asp	Glu	Pro	Phe	Ser	Leu	Gln	Ala	Met	Lys	Glu	275	280	285	
Ala	Ala	Thr	Glu	Leu	Leu	Phe	Gly	Gly	His	Glu	Thr	Thr	Ala	Ser	Thr	290	295	300	
Ala	Thr	Ser	Leu	Val	Met	Phe	Leu	Gly	Leu	Asn	Thr	Glu	Val	Val	Gln				



305		310		315		320
Lys Val Arg Glu	Glu Val Gln Glu Lys	Val Glu Met Gly Met Tyr Thr				
	325	330			335	
Pro Gly Lys Gly	Leu Ser Met Glu Leu Leu Asp Gln Leu Lys Tyr Thr					
	340	345			350	
Gly Cys Val Ile	Lys Glu Thr Leu Arg Ile Asn Pro Pro Val Pro Gly					
	355	360			365	
Gly Phe Arg Val	Ala Leu Lys Thr Phe Glu Leu Asn Gly Tyr Gln Ile					
	370	375			380	
Pro Lys Gly Trp	Asn Val Ile Tyr Ser Ile Cys Asp Thr His Asp Val					
	385	390			395	400
Ala Asp Val Phe	Pro Asn Lys Glu Glu Phe Gln Pro Glu Arg Phe Met					
	405	410			415	
Ser Lys Gly Leu	Glu Asp Gly Ser Arg Phe Asn Tyr Ile Pro Phe Gly					
	420	425			430	
Gly Gly Ser Arg	Met Cys Val Gly Lys Glu Phe Ala Lys Val Leu Leu					
	435	440			445	
Lys Ile Phe Leu	Val Glu Leu Thr Gln His Cys Asn Trp Ile Leu Ser					
	450	455			460	
Asn Gly Pro Pro	Thr Met Lys Thr Gly Pro Thr Ile Tyr Pro Val Asp					
	465	470			475	480
Asn Leu Pro Thr	Lys Phe Thr Ser Tyr Val Arg Asn					
	485	490				

(2) INFORMATION FOR SEQ ID NO:3

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1850 base pairs
- (B) TYPE: nucleic acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3

TGTCGCCGTT GCTGTCGGTT GCTGTCCGAC GCTGTCTCCT CTCCAGAAGC TTGTTTTTCG	60
TTTTGGCGAT CAGTTGCGCG CTTCAAC ATG GGG CTG TAC ACC CTT ATG GTC ACC	114
Met Gly Leu Tyr Thr Leu Met Val Thr	
1 5	
TTT CTC TGC ACC ATC GTG CTA CCC GTT TTA CTC TTT CTC GCC GCG GTG	162
Phe Leu Cys Thr Ile Val Leu Pro Val Leu Leu Phe Leu Ala Ala Val	
10 15 20 25	
AAG TTG TGG GAG ATG TTA ATG ATC CGA CGA GTC GAT CCG AAC TGC AGA	210
Lys Leu Trp Glu Met Leu Met Ile Arg Arg Val Asp Pro Asn Cys Arg	
30 35 40	
AGT CCT CTA CCG CCA GGT ACC ATG GGC TTG CCG TTC ATT GGA GAA ACG	258
Ser Pro Leu Pro Pro Gly Thr Met Gly Leu Pro Phe Ile Gly Glu Thr	
45 50 55	
CTC CAG CTG ATC CTC CAG AGA AGG AAG TTT CTG CGC ATG AAA CGG CAG	306
Leu Gln Leu Ile Leu Gln Arg Arg Lys Phe Leu Arg Met Lys Arg Gln	
60 65 70	

00568482 092500



AAA TAC GGG TGC ATC TAC AAG ACG CAC CTC TTC GGG AAC CCG ACT GTC	354
Lys Tyr Gly Cys Ile Tyr Lys Thr His Leu Phe Gly Asn Pro Thr Val	
75 80 85	
AGG GTG ATG GGA GCT GAT AAT GTG AGG CAG ATT CTG CTG GGC GAA CAC	402
Arg Val Met Gly Ala Asp Asn Val Arg Gln Ile Leu Leu Gly Glu His	
90 95 100 105	
AAG CTG GTG TCT GTT CAG TGG CCA GCA TCA GTG AGA ACC ATC CTG GGC	450
Lys Leu Val Ser Val Gln Trp Pro Ala Ser Val Arg Thr Ile Leu Gly	
110 115 120	
TCT GAC ACC CTC TCC AAT GTC CAT GGA GTT CAA CAC AAA AAC AAG AAA	498
Ser Asp Thr Leu Ser Asn Val His Gly Val Gln His Lys Asn Lys Lys	
125 130 135	
AAG GCC ATT ATG AGG GCG TTC TCT CGA GAT GCT CTG GAG CAC TAC ATT	546
Lys Ala Ile Met Arg Ala Phe Ser Arg Asp Ala Leu Glu His Tyr Ile	
140 145 150	
CCC GTG ATC CAG CAG GAG GTG AAG AGC GCC ATA CAG GAA TGG CTG CAA	594
Pro Val Ile Gln Gln Glu Val Lys Ser Ala Ile Gln Glu Trp Leu Gln	
155 160 165	
AAA GAC TCC TGC GTG CTG GTT TAT CCA GAA ATG AAG AAA CTC ATG TTT	642
Lys Asp Ser Cys Val Leu Val Tyr Pro Glu Met Lys Lys Leu Met Phe	
170 175 180 185	
CGG ATA GCT ATG AGA ATC CTG CTT GGT TTT GAA CCA GAG CAA ATA AAG	690
Arg Ile Ala Met Arg Ile Leu Leu Gly Phe Glu Pro Glu Gln Ile Lys	
190 195 200	
ACG GAC GAG CAA GAA CTG GTG GAA GCT TTT GAG GAA ATG ATC AAA AAC	738
Thr Asp Glu Gln Glu Leu Val Glu Ala Phe Glu Glu Met Ile Lys Asn	
205 210 215	
TTG TTC TCC TTG CCA ATC GAC GTT CCT TTC AGT GGT CTG TAC AGG GGT	786
Leu Phe Ser Leu Pro Ile Asp Val Pro Phe Ser Gly Leu Tyr Arg Gly	
220 225 230	
TTG AGG GCA CGC AAT TTC ATT CAC TCC AAA ATT GAG GAA AAC ATC AGG	834
Leu Arg Ala Arg Asn Phe Ile His Ser Lys Ile Glu Glu Asn Ile Arg	
235 240 245	
AAG AAA ATT CAA GAT GAC GAC AAT GAA AAC GAA CAG AAA TAC AAA GAC	882
Lys Lys Ile Gln Asp Asp Asp Asn Glu Asn Glu gln Lys Tyr Lys Asp	
250 255 260 265	
GCC CTT CAG CTG TTG ATC GAG AAC AGC AGA AGA AGT GAC GAA CCT TTT	930
Ala Leu Gln Leu Leu Ile Glu Asn Ser Arg Arg Ser Asp Glu Pro Phe	
270 275 280	
AGT TTG CAG GCG ATG AAA GAA GCA GCT ACA GAG CTT CTA TTT GGA GGT	978
Ser Leu Gln Ala Met Lys Glu Ala Ala Thr Glu Leu Leu Phe Gly Gly	
285 290 295	
CAT GAA ACC ACC GCC AGC ACT GCA ACC TCA CTT GTC ATG TTT CTG GGT	1026
His Glu Thr Thr Ala Ser Thr Ala Thr Ser Leu Val Met Phe Leu Gly	
300 305 310	
CTG AAC ACA GAA GTG GTG CAG AAG GTC AGA GAG GAG GTT CAG GAG AAG	1074
Leu Asn Thr Glu Val Val Gln Lys Val Arg Glu Glu Val Gln Glu Lys	
315 320 325	

005250" 28489950



GTT GAA ATG GGC ATG TAT ACA CCT GGA AAG GGC TTG AGT ATG GAG CTG	1122
Val Glu Met Gly Met Tyr Thr Pro Gly Lys Gly Leu Ser Met Glu Leu	
330 335 340 345	
TTG GAC CAG CTG AAG TAC ACT GGA TGT GTG ATT AAA GAG ACT CTT AGA	1170
Leu Asp Gln Leu Lys Tyr Thr Gly Cys Val Ile Lys Glu Thr Leu Arg	
350 355 360	
ATC AAC CCT CCT GTT CCC GGA GGA TTC AGA GTC GCA CTC AAA ACC TTT	1218
Ile Asn Pro Pro Val Pro Gly Gly Phe Arg Val Ala Leu Lys Thr Phe	
365 370 375	
GAA TTG AAT GGT TAC CAA ATT CCT AAA GGA TGG AAC GTC ATT TAC AGC	1266
Glu Leu Asn Gly Tyr Gln Ile Pro Lys Gly Trp Asn Val Ile Tyr Ser	
380 385 390	
ATC TGT GAC ACG CAC GAT GTG GCC GAC GTC TTT CCA AAC AAA GAG GAG	1314
Ile Cys Asp Thr His Asp Val Ala Asp Val Phe Pro Asn Lys Glu Glu	
395 400 405	
TTC CAG CCG GAG AGA TTC ATG AGC AAA GGT CTG GAG GAC GGG TCC AGG	1362
Phe Gln Pro Glu Arg Phe Met Ser Lys Gly Leu Glu Asp Gly Ser Arg	
410 415 420 425	
TTT AAC TAC ATC CCC TTC GGA GGA GGA TCC AGG ATG TGT GTG GGC AAA	1410
Phe Asn Tyr Ile Pro Phe Gly Gly Gly Ser Arg Met Cys Val Gly Lys	
430 435 440	
GAG TTC GCC AAA GTG TTA CTC AAG ATC TTT TTA GTT GAG TTA ACG CAG	1458
Glu Phe Ala Lys Val Leu Leu Lys Ile Phe Leu Val Glu Leu Thr Gln	
445 450 455	
CAT TGC AAT TGG ATT CTC TCA AAC GGA CCC CCG ACA ATG AAA ACA GGC	1506
His Cys Asn Trp Ile Leu Ser Asn Gly Pro Pro Thr Met Lys Thr Gly	
460 465 470	
CCG ACT ATT TAC CCA GTG GAC AAT CTC CCT ACC AAA TTC ACT AGT TAT	1554
Pro Thr Ile Tyr Pro Val Asp Asn Leu Pro Thr Lys Phe Thr Ser Tyr	
475 480 485	
GTC AGA AAT TAGCCTAACC GGAGCTTTGT ACATATGTTT TTATTTTAGA	1603
Val Arg Asn	
490	
TGAACTGTGA TGTATTGGAT ATTTTCTATT TTGTTTATAT AAAGCAGATG TGTATATAAG	1663
TCTATGCGAG GAAGCGAAAA CGAGGGCACT ACTTTCTCAT GGATCACTGT AATGCTACAG	1723
AGTGTCTGTG ATGTATATTT ATAATGTAGT TGTGTTATAT AGCTTTTGTA CTGTATGCAA	1783
CTTATTTAAC TCGCTCTTTA TCTCATGGGT TTTATTTAAT AAAACATGTT CTTACAAAAA	1843
AAAAAAA	1850

(2) INFORMATION FOR SEQ ID NO:4

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 497 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4

Met	Gly	Leu	Pro	Ala	Leu	Leu	Ala	Ser	Ala	Leu	Cys	Thr	Phe	Val	Leu
1					5				10					15	



Pro Leu Leu Leu Phe Leu Ala Ala Ile Lys Leu Trp Asp Leu Tyr Cys  
20 25 30

Val Ser Gly Arg Asp Arg Ser Cys Ala Leu Pro Leu Pro Gly Thr  
35 40 45

Met Gly Phe Pro Phe Phe Gly Glu Thr Leu Gln Met Val Leu Gln Arg  
50 55 60

Arg Lys Phe Leu Gln Met Lys Arg Arg Lys Tyr Gly Phe Ile Tyr Lys  
65 70 75 80

Thr His Leu Phe Gly Arg Pro Thr Val Arg Val Met Gly Ala Asp Asn  
85 90 95

Val Arg Arg Ile Leu Leu Gly Asp Asp Arg Leu Val Ser Val His Trp  
100 105 110

Pro Ala Ser Val Arg Thr Ile Leu Gly Ser Gly Cys Leu Ser Asn Leu  
115 120 125

His Asp Ser Ser His Lys Gln Arg Lys Lys Val Ile Met Arg Ala Phe  
130 135 140

Ser Arg Glu Ala Leu Glu Cys Tyr Val Pro Val Ile Thr Glu Glu Val  
145 150 155 160

Gly Ser Ser Leu Glu Gln Trp Leu Ser Cys Gly Glu Arg Gly Leu Leu  
165 170 175

Val Tyr Pro Glu Val Lys Arg Leu Met Phe Arg Ile Ala Met Arg Ile  
180 185 190

Leu Leu Gly Cys Glu Pro Gln Leu Ala Gly Asp Gly Asp Ser Glu Gln  
195 200 205

Gln Leu Val Glu Ala Phe Glu Glu Met Thr Arg Asn Leu Phe Ser Leu  
210 215 220

Pro Ile Asp Val Pro Phe Ser Gly Leu Tyr Arg Gly Met Lys Ala Arg  
225 230 235 240

Asn Leu Ile His Ala Arg Ile Glu Gln Asn Ile Arg Ala Lys Ile Cys  
245 250 255

Gly Leu Arg Ala Ser Glu Ala Gly Gln Gly Cys Lys Asp Ala Leu Gln  
260 265 270

Leu Leu Ile Glu His Ser Trp Glu Arg Gly Glu Arg Leu Asp Met Gln  
275 280 285

Ala Leu Lys Gln Ser Ser Thr Glu Leu Leu Phe Gly Gly His Glu Thr  
290 295 300

Thr Ala Ser Ala Ala Thr Ser Leu Ile Thr Tyr Leu Gly Leu Tyr Pro  
305 310 315 320

His Val Leu Gln Lys Val Arg Glu Glu Leu Lys Ser Lys Gly Leu Leu  
325 330 335

Cys Lys Ser Asn Gln Asp Asn Lys Leu Asp Met Glu Ile Leu Glu Gln  
340 345 350

Leu Lys Tyr Ile Gly Cys Val Ile Lys Glu Thr Leu Arg Leu Asn Pro  
355 360 365

09668482 "092500"



Pro Val Pro Gly Gly Phe Arg Val Ala Leu Lys Thr Phe Glu Leu Asn  
370 375 380

Gly Tyr Gln Ile Pro Lys Gly Trp Asn Val Ile Tyr Ser Ile Cys Asp  
385 390 395 400

Thr His Asp Val Ala Glu Ile Phe Thr Asn Lys Glu Glu Phe Asn Pro  
405 410 415

Asp Arg Phe Ser Ala Pro His Pro Glu Asp Ala Ser Arg Phe Ser Phe  
420 425 430

Ile Pro Phe Gly Gly Gly Leu Arg Ser Cys Val Gly Lys Glu Phe Ala  
435 440 445

Lys Ile Leu Leu Lys Ile Phe Thr Val Glu Leu Ala Arg His Cys Asp  
450 455 460

Trp Gln Leu Leu Asn Gly Pro Pro Thr Met Lys Thr Ser Pro Thr Val  
465 470 475 480

Tyr Pro Val Asp Asn Leu Pro Ala Arg Phe Thr His Phe His Gly Glu  
485 490 495

Ile

(2) INFORMATION FOR SEQ ID NO:5

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1494 base pairs
- (B) TYPE: nucleic acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5

ATG GGG CTC CCG GCG CTG CTG GCC AGT GCG CTC TGC ACC TTC GTG CTG	48
Met Gly Leu Pro Ala Leu Leu Ala Ser Ala Leu Cys Thr Phe Val Leu	
1 5 10 15	
CCG CTG CTG CTC TTC CTG GCT GCG ATC AAG CTC TGG GAC CTG TAC TGC	96
Pro Leu Leu Leu Phe Leu Ala Ala Ile Lys Leu Trp Asp Leu Tyr Cys	
20 25 30	
GTG AGC GGC CGC GAC CGC AGT TGT GCC CTC CCA TTG CCC CCC GGG ACT	144
Val Ser Gly Arg Asp Arg Ser Cys Ala Leu Pro Leu Pro Pro Gly Thr	
35 40 45	
ATG GGC TTC CCC TTC TTT GGG GAA ACC TTG CAG ATG GTA CTG CAG CGG	192
Met Gly Phe Pro Phe Phe Gly Glu Thr Leu Gln Met Val Leu Gln Arg	
50 55 60	
AGG AAG TTC CTG CAG ATG AAG CGC AGG AAA TAC GGC TTC ATC TAC AAG	240
Arg Lys Phe Leu Gln Met Lys Arg Arg Lys Tyr Gly Phe Ile Tyr Lys	
65 70 75 80	
ACG CAT CTG TTC GGG CGG CCC ACC GTA CGG GTG ATG GGC GCG GAC AAT	288
Thr His Leu Phe Gly Arg Pro Thr Val Arg Val Met Gly Ala Asp Asn	
85 90 95	
GTG CGG CGC ATC TTG CTC GGA GAC GAC CGG CTG GTG TCG GTC CAC TGG	336
Val Arg Arg Ile Leu Leu Gly Asp Asp Arg Leu Val Ser Val His Trp	
100 105 110	

09663487 092500 005260 "28489960"



CCA	GCG	TCG	GTG	CGC	ACC	ATT	CTG	GGA	TCT	GGC	TGC	CTC	TCT	AAC	CTG	384
Pro	Ala	Ser	Val	Arg	Thr	Ile	Leu	Gly	Ser	Gly	Cys	Leu	Ser	Asn	Leu	
		115					120					125				
CAC	GAC	TCC	TCG	CAC	AAG	CAG	CGC	AAG	AAG	GTG	ATT	ATG	CGG	GCC	TTG	432
His	Asp	Ser	Ser	His	Lys	Gln	Arg	Lys	Lys	Val	Ile	Met	Arg	Ala	Phe	
		130				135					140					
AGC	CGC	GAG	GCA	CTC	GAA	TGC	TAC	GTG	CCG	GTG	ATC	ACC	GAG	GAA	GTG	480
Ser	Arg	Glu	Ala	Leu	Glu	Cys	Tyr	Val	Pro	Val	Ile	Thr	Glu	Glu	Val	
		145				150				155					160	
GGC	AGC	AGC	CTG	GAG	CAG	TGG	CTG	AGC	TGC	GGC	GAG	CGC	GGC	CTC	CTG	528
Gly	Ser	Ser	Leu	Glu	Gln	Trp	Leu	Ser	Cys	Gly	Glu	Arg	Gly	Leu	Leu	
				165					170					175		
GTC	TAC	CCC	GAG	GTG	AAG	CGC	CTC	ATG	TTC	CGA	ATC	GCC	ATG	CGC	ATC	576
Val	Tyr	Pro	Glu	Val	Lys	Arg	Leu	Met	Phe	Arg	Ile	Ala	Met	Arg	Ile	
			180					185					190			
CTA	CTG	GGC	TGC	GAA	CCC	CAA	CTG	GCG	GGC	GAC	GGG	GAC	TCC	GAG	CAG	624
Leu	Leu	Gly	Cys	Glu	Pro	Gln	Leu	Ala	Gly	Asp	Gly	Asp	Ser	Glu	Gln	
		195					200					205				
CAG	CTT	GTG	GAG	GCC	TTC	GAG	GAA	ATG	ACC	CGC	AAT	CTC	TTC	TCG	CTG	672
Gln	Leu	Val	Glu	Ala	Phe	Glu	Glu	Met	Thr	Arg	Asn	Leu	Phe	Ser	Leu	
		210				215					220					
CCC	ATC	GAC	GTG	CCC	TTC	AGC	GGG	CTG	TAC	CGG	GGC	ATG	AAG	GCG	CGG	720
Pro	Ile	Asp	Val	Pro	Phe	Ser	Gly	Leu	Tyr	Arg	Gly	Met	Lys	Ala	Arg	
		225			230					235					240	
AAC	CTC	ATT	CAC	GCG	CGC	ATC	GAG	CAG	AAC	ATT	CGC	GCC	AAG	ATC	TGC	768
Asn	Leu	Ile	His	Ala	Arg	Ile	Glu	Gln	Asn	Ile	Arg	Ala	Lys	Ile	Cys	
				245					250					255		
GGG	CTG	CGG	GCA	TCC	GAG	GCG	GGC	CAG	GGC	TGC	AAA	GAC	GCG	CTG	CAG	816
Gly	Leu	Arg	Ala	Ser	Glu	Ala	Gly	Gln	Gly	Cys	Lys	Asp	Ala	Leu	Gln	
			260					265					270			
CTG	TTG	ATC	GAG	CAC	TCG	TGG	GAG	AGG	GGA	GAG	CGG	CTG	GAC	ATG	CAG	864
Leu	Leu	Ile	Glu	His	Ser	Trp	Glu	Arg	Gly	Glu	Arg	Leu	Asp	Met	Gln	
		275					280					285				
GCA	CTA	AAG	CAA	TCT	TCA	ACC	GAA	CTC	CTC	TTT	GGA	GGA	CAC	GAA	ACC	912
Ala	Leu	Lys	Gln	Ser	Ser	Thr	Glu	Leu	Leu	Phe	Gly	Gly	His	Glu	Thr	
		290				295					300					
ACG	GCC	AGT	GCA	GCC	ACA	TCT	CTG	ATC	ACT	TAC	CTG	GGG	CTC	TAC	CCA	960
Thr	Ala	Ser	Ala	Ala	Thr	Ser	Leu	Ile	Thr	Tyr	Leu	Gly	Leu	Tyr	Pro	
		305			310					315					320	
CAT	GTT	CTC	CAG	AAA	GTG	CGA	GAA	GAG	CTG	AAG	AGT	AAG	GGT	TTA	CTT	1008
His	Val	Leu	Gln	Lys	Val	Arg	Glu	Glu	Leu	Lys	Ser	Lys	Gly	Leu	Leu	
				325					330					335		
TGC	AAG	AGC	AAT	CAA	GAC	AAC	AAG	TTG	GAC	ATG	GAA	ATT	TTG	GAA	CAA	1056
Cys	Lys	Ser	Asn	Gln	Asp	Asn	Lys	Leu	Asp	Met	Glu	Ile		Glu	Gln	
			340					345					350			
CTT	AAA	TAC	ATC	GGG	TGT	GTT	ATT	AAG	GAG	ACC	CTT	CGA	CTG	AAT	CCC	1104
Leu	Lys	Tyr	Ile	Gly	Cys	Val	Ile	Lys	Glu	Thr	Leu	Arg	Leu	Asn	Pro	
		355					360					365				

005260-22429960



CCA GTT CCA GGA GGG TTT CGG GTT GCT CTG AAG ACT TTT GAA TTA AAT	1152
Pro Val Pro Gly Gly Phe Arg Val Ala Leu Lys Thr Phe Glu Leu Asn	
370 375 380	
GGA TAC CAG ATT CCC AAG GGC TGG AAT GTT ATC TAC AGT ATC TGT GAT	1200
Gly Tyr Gln Ile Pro Lys Gly Trp Asn Val Ile Tyr Ser Ile Cys Asp	
385 390 395 400	
ACT CAT GAT GTG GCA GAG ATC TTC ACC AAC AAG GAA GAA TTT AAT CCT	1248
Thr His Asp Val Ala Glu Ile Phe Thr Asn Lys Glu Glu Phe Asn Pro	
405 410 415	
GAC CGA TTC AGT GCT CCT CAC CCA GAG GAT GCA TCC AGG TTC AGC TTC	1296
Asp Arg Phe Ser Ala Pro His Pro Glu Asp Ala Ser Arg Phe Ser Phe	
420 425 430	
ATT CCA TTT GGA GGA GGC CTT AGG AGC TGT GTA GGC AAA GAA TTT GCA	1344
Ile Pro Phe Gly Gly Gly Leu Arg Ser Cys Val Gly Lys Glu Phe Ala	
435 440 445	
AAA ATT CTT CTC AAA ATA TTT ACA GTG GAG CTG GCC AGG CAT TGT GAC	1392
Lys Ile Leu Leu Lys Ile Phe Thr Val Glu Leu Ala Arg His Cys Asp	
450 455 460	
TGG CAG CTT CTA AAT GGA CCT CCT ACA ATG AAA ACC AGT CCC ACC GTG	1440
Trp Gln Leu Leu Asn Gly Pro Pro Thr Met Lys Thr Ser Pro Thr Val	
465 470 475 480	
TAT CCT GTG GAC AAT CTC CCT GCA AGA TTC ACC CAT TTC CAT GGG GAA	1488
Tyr Pro Val Asp Asn Leu Pro Ala Arg Phe Thr His Phe His Gly Glu	
485 490 495	
ATC TGA	1494
Ile	

(2) INFORMATION FOR SEQ ID NO:6

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6

Pro Phe Gly Gly Gly Pro Arg Leu Cys	Pro Gly Tyr Glu Leu Ala Arg
1 5 10 15	
Val Ala Leu Ser	
20	

(2) INFORMATION FOR SEQ ID NO:7

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7

Pro Phe Ser Gly Gly Ala Arg Asn Cys	Ile Gly Lys Gln Phe Ala Met
1 5 10 15	
Ser Glu Met Lys	
20	

005260" 28489960



(2) INFORMATION FOR SEQ ID NO:8

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8

Pro Phe Ser Gly Gly Ala Arg Asn Cys Ile Gly Lys Gln Phe Ala Met  
1 5 10 15  
Asn Glu Leu Lys  
20

(2) INFORMATION FOR SEQ ID NO:9

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9

Pro Phe Gly Thr Gly Pro Arg Asn Cys Ile Gly Met Arg Phe Ala Ile  
1 5 10 15  
Met Asn Met Lys  
20

(2) INFORMATION FOR SEQ ID NO:10

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10

Pro Phe Ser Gly Gly Ser Arg Asn Cys Ile Gly Lys Gln Phe Ala Met  
1 5 10 15  
Asn Glu Leu Lys  
20

(2) INFORMATION FOR SEQ ID NO:11

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 351 base pairs  
(B) TYPE: nucleic acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11

GAACTCCTCT TTGGAGGACA CGAAACCACG GCCAGTGCAG CCACATCTCT GATCACTTAC	60
CTGGGGCTCT ACCCACATGT TCTCCAGAAA GTGCGAGAAG AGCTGAAGAG TAAGGGTTTA	120
CTTTGCAAGA GCAATCAAGA CAACAAGTTG GACATGGAAA TTTTGGAACA ACTTAAATAC	180
ATCGGGTGTG TTATTAAGGA GACCCTTCGA CTGAATCCCC CAGTTCCAGG AGGGTTTCGG	240
GTGCTCTGA AGACTTTTGA ATTAAATGGA TACCAGATTC CCAAGGGCTG GAATGTTATC	300
TACAGTATCT GTGATACTCA TGATGTGGCA GAGATCTTCA CCAACAAGGA A	351



- (2) INFORMATION FOR SEQ ID NO:12  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 14 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12

TTTTTTTTTT TTGG

14

- (2) INFORMATION FOR SEQ ID NO:13  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 14 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13

TTTTTTTTTT TTGA

14

- (2) INFORMATION FOR SEQ ID NO:14  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 14 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14

TTTTTTTTTT TTGT

14

- (2) INFORMATION FOR SEQ ID NO:15  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 14 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15

TTTTTTTTTT TTGC

14

- (2) INFORMATION FOR SEQ ID NO:16  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 14 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16

TTTTTTTTTT TTAG

14

- (2) INFORMATION FOR SEQ ID NO:17  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 14 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17

TTTTTTTTTT TTAA

14

005260"28489360



- (2) INFORMATION FOR SEQ ID NO:18  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 14 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18

TTTTTTTTTT TTAT

14

- (2) INFORMATION FOR SEQ ID NO:19  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 14 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19

TTTTTTTTTT TTAC

14

- (2) INFORMATION FOR SEQ ID NO:20  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 14 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20

TTTTTTTTTT TTCG

14

- (2) INFORMATION FOR SEQ ID NO:21  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 14 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21

TTTTTTTTTT TTCA

14

- (2) INFORMATION FOR SEQ ID NO:22  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 14 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22

TTTTTTTTTT TTCT

14

- (2) INFORMATION FOR SEQ ID NO:23  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 14 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23

TTTTTTTTTT TTCC

14

0968482 092500



- (2) INFORMATION FOR SEQ ID NO:24  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 10 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24

AAGCGACCGA

10

- (2) INFORMATION FOR SEQ ID NO:25  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 10 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25

TGTTCCGCCAG

10

- (2) INFORMATION FOR SEQ ID NO:26  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 10 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26

TGCCAGTGGA

10

- (2) INFORMATION FOR SEQ ID NO:27  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 10 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27

GGCTGCAAAC

10

- (2) INFORMATION FOR SEQ ID NO:28  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 10 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28

CCTAGCGTTG

10

- (2) INFORMATION FOR SEQ ID NO:29  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 21 base pairs  
        (B) TYPE: nucleic acid  
        (D) TOPOLOGY: linear  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29

GTAGCGGCCG CTGCCAGTGG A

21



(2) INFORMATION FOR SEQ ID NO:30  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 12 base pairs  
 (B) TYPE: nucleic acid  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30

GTAGCGGCCG CT

12

(2) INFORMATION FOR SEQ ID NO:31  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1725 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31 .

GCACGAGGGA GGCTGAAGCG TGCC ATG GGG CTC CCG GCG CTG CTG GCC AGT	51
Met Gly Leu Pro Ala Leu Leu Ala Ser	
1 5	
GCG CTC TGC ACC TTC GTG CTG CCG CTG CTG CTC TTC CTG GCG GCG CTC	99
Ala Leu Cys Thr Phe Val Leu Pro Leu Leu Leu Phe Leu Ala Ala Leu	
10 15 20 25	
AAG CTC TGG GAC CTG TAC TGT GTG AGC AGC CGC GAT CGC AGC TGC GCC	147
Lys Leu Trp Asp Leu Tyr Cys Val Ser Ser Arg Asp Arg Ser Cys Ala	
30 35 40	
CTC CCC TTG CCC CCC GGT ACC ATG GGC TTC CCA TTC TTT GGG GAA ACA	195
Leu Pro Leu Pro Pro Gly Thr Met Gly Phe Pro Phe Phe Gly Glu Thr	
45 50 55	
TTG CAG ATG GTG CTT CAG CGG AGG AAG TTT CTG CAG ATG AAG CGC AGG	243
Leu Gln Met Val Leu Gln Arg Arg Lys Phe Leu Gln Met Lys Arg Arg	
60 65 70	
AAA TAC GGC TTC ATC TAC AAG ACG CAT CTG TTT GGG CGG CCC ACG GTG	291
Lys Tyr Gly Phe Ile Tyr Lys Thr His Leu Phe Gly Arg Pro Thr Val	
75 80 85	
CGG GTG ATG GGC GCG GAT AAT GTG CGG CGC ATC TTG CTG GGA GAG CAC	339
Arg Val Met Gly Ala Asp Asn Val Arg Arg Ile Leu Leu Gly Glu His	
90 95 100 105	
CGG TTG GTG TCG GTG CAC TGG CCC GCG TCG GTG CGC ACC ATC CTG GGC	387
Arg Leu Val Ser Val His Trp Pro Ala Ser Val Arg Thr Ile Leu Gly	
110 115 120	
GCT GGC TGC CTC TCC AAC CTG CAC GAT TCC TCG CAC AAG CAG CGA AAG	435
Ala Gly Cys Leu Ser Asn Leu His Asp Ser Ser His Lys Gln Arg Lys	
125 130 135	
AAG GTG ATT ATG CAG GCC TTC AGC CGC GAG GCA CTC CAG TGC TAC GTG	483
Lys Val Ile Met Gln Ala Phe Ser Arg Glu Ala Leu Gln Cys Tyr Val	
140 145 150	
CTC GTG ATC GCT GAG GAA GTC AGC AGT TGT CTG GAG CAG TGG CTA AGC	531
Leu Val Ile Ala Glu Glu Val Ser Ser Cys Leu Glu Gln Trp Leu Ser	
155 160 165	

005260 "2848960



TGC GGC GAG CGC GGC CTC CTG GTC TAC CCC GAG GTG AAG CGC CTC ATG Cys Gly Glu Arg Gly Leu Leu Val Tyr Pro Glu Val Lys Arg Leu Met 170 175 180 185	579
TTC CGC ATC GCC ATG CGC ATC CTG CTG GGC TGC GAG CCG GGT CCA GCG Phe Arg Ile Ala Met Arg Ile Leu Leu Gly Cys Glu Pro Gly Pro Ala 190 195 200	627
GGC GGC GGG GAG GAC GAG CAA CAG CTC GTG GAG GCT TTC GAG GAG ATG Gly Gly Gly Glu Asp Glu Gln Gln Leu Val Glu Ala Phe Glu Glu Met 205 210 215	675
ACC CGC AAT CTC TTC TCT CTT CCC ATT GAC GTG CCC TTT AGC GGC CTG Thr Arg Asn Leu Phe Ser Leu Pro Ile Asp Val Pro Phe Ser Gly Leu 220 225 230	723
TAC CGG GGC GTG AAG GCG CGG AAC CTT ATA CAC GCG CGC ATC GAG GAG Tyr Arg Gly Val Lys Ala Arg Asn Leu Ile His Ala Arg Ile Glu Glu 235 240 245	771
AAC ATT CGC GCC AAG ATC CGC CGG CTT CAG GCT ACA GAG CCG GAT GGG Asn Ile Arg Ala Lys Ile Arg Arg Leu Gln Ala Thr Glu Pro Asp Gly 250 255 260 265	819
GGT TGC AAG GAC GCG CTG CAG CTC CTG ATT GAG CAC TCG TGG GAG AGG Gly Cys Lys Asp Ala Leu Gln Leu Leu Ile Glu His Ser Trp Glu Arg 270 275 280	867
GGA GAG AGG CTG GAT ATG CAG GCA CTA AAA CAA TCG TCA ACA GAG CTC Gly Glu Arg Leu Asp Met Gln Ala Leu Lys Gln Ser Ser Thr Glu Leu 285 290 295	915
CTC TTT GGT GGT CAT GAA ACT ACA GCC AGT GCT GCG ACA TCA CTG ATC Leu Phe Gly Gly His Glu Thr Thr Ala Ser Ala Ala Thr Ser Leu Ile 300 305 310	963
ACT TAC CTA GGA CTC TAC CCA CAT GTC CTC CAG AAA GTT CGA GAA GAG Thr Tyr Leu Gly Leu Tyr Pro His Val Leu Gln Lys Val Arg Glu Glu 315 320 325	1011
ATA AAG AGC AAG GGC TTA CTT TGC AAG AGC AAT CAA GAC AAC AAG TTA Ile Lys Ser Lys Gly Leu Leu Cys Lys Ser Asn Gln Asp Asn Lys Leu 330 335 340 345	1059
GAC ATG GAA ACT TTG GAA CAG CTT AAA TAC ATT GGG TGT GTC ATT AAG Asp Met Glu Thr Leu Glu Gln Leu Lys Tyr Ile Gly Cys Val Ile Lys 350 355 360	1107
GAG ACC CTG CGA TTG AAT CCT CCG GTT CCA GGA GGG TTT CGG GTT GCT Glu Thr Leu Arg Leu Asn Pro Pro Val Pro Gly Gly Phe Arg Val Ala 365 370 375	1155
CTG AAG ACT TTT GAG CTG AAT GGA TAC CAG ATC CCC AAG GGC TGG AAT Leu Lys Thr Phe Glu Leu Asn Gly Tyr Gln Ile Pro Lys Gly Trp Asn 380 385 390	1203
GTT ATT TAC AGT ATC TGT GAC ACC CAC GAT GTG GCA GAT ATC TTC ACT Val Ile Tyr Ser Ile Cys Asp Thr His Asp Val Ala Asp Ile Phe Thr 395 400 405	1251
AAC AAG GAG GAA TTT AAT CCC GAC CGC TTT ATA GTG CCT CAT CCA GAG Asn Lys Glu Glu Phe Asn Pro Asp Arg Phe Ile Val Pro His Pro Glu 410 415 420 425	1299

09560482 092500



GAT GCT TCC CGG TTC AGC TTC ATT CCA TTT GGA GGA GGC CTT CGG AGC	1347
Asp Ala Ser Arg Phe Ser Phe Ile Pro Phe Gly Gly Gly Leu Arg Ser	
430 435 440	
TGT GTA GGC AAA GAG TTT GCA AAA ATT CTT CTT AAG ATA TTT ACA GTG	1395
Cys Val Gly Lys Glu Phe Ala Lys Ile Leu Leu Lys Ile Phe Thr Val	
445 450 455	
GAG CTG GCT AGG CAC TGT GAT TGG CAG CTT CTA AAT GGA CCT CCT ACA	1443
Glu Leu Ala Arg His Cys Asp Trp Gln Leu Leu Asn Gly Pro Pro Thr	
460 465 470	
ATG AAG ACA AGC CCC ACT GTG TAC CCT GTG GAC AAT CTC CCT GCA AGA	1491
Met Lys Thr Ser Pro Thr Val Tyr Pro Val Asp Asn Leu Pro Ala Arg	
475 480 485	
TTC ACC TAC TTC CAG GGA GAT ATC TGATAGCTAT TTCAATTCTT	1535
Phe Thr Tyr Phe Gln Gly Asp Ile	
490 495	
GGACTTATTT GAAGTGATA TTGGTTTTTTT TAAAAATAG TGTCATGTTG ACTTTATTTA	1595
ATTTCTAAAT GTATAGTATG ATATTTATGT GTCTCTACTA CAGTCCCGTG GTCTTTAAAT	1655
ATTAAAATAA TGAATTTGTA TGATTTCCCA ATAAAGTAAA ATTAAAAAGT GAAAAAATAA	1715
AAAAAATAA	1725

(2) INFORMATION FOR SEQ ID NO:32

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 497 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32

Met Gly Leu Pro Ala Leu Leu Ala Ser Ala Leu Cys Thr Phe Val Leu	
1 5 10 15	
Pro Leu Leu Leu Phe Leu Ala Ala Leu Lys Leu Trp Asp Leu Tyr Cys	
20 25 30	
Val Ser Ser Arg Asp Arg Ser Cys Ala Leu Pro Leu Pro Pro Gly Thr	
35 40 45	
Met Gly Phe Pro Phe Phe Gly Glu Thr Leu Gln Met Val Leu Gln Arg	
50 55 60	
Arg Lys Phe Leu Gln Met Lys Arg Arg Lys Tyr Gly Phe Ile Tyr Lys	
65 70 75 80	
Thr His Leu Phe Gly Arg Pro Thr Val Arg Val Met Gly Ala Asp Asn	
85 90 95	
Val Arg Arg Ile Leu Leu Gly Glu His Arg Leu Val Ser Val His Trp	
100 105 110	
Pro Ala Ser Val Arg Thr Ile Leu Gly Ala Gly Cys Leu Ser Asn Leu	
115 120 125	
His Asp Ser Ser His Lys Gln Arg Lys Lys Val Ile Met Gln Ala Phe	
130 135 140	

005260" 2842960



Ser	Arg	Glu	Ala	Leu	Gln	Cys	Tyr	Val	Leu	Val	Ile	Ala	Glu	Glu	Val
145					150					155					160
Ser	Ser	Cys	Leu	Glu	Gln	Trp	Leu	Ser	Cys	Gly	Glu	Arg	Gly	Leu	Leu
			165						170					175	
Val	Tyr	Pro	Glu	Val	Lys	Arg	Leu	Met	Phe	Arg	Ile	Ala	Met	Arg	Ile
			180					185					190		
Leu	Leu	Gly	Cys	Glu	Pro	Gly	Pro	Ala	Gly	Gly	Gly	Glu	Asp	Glu	Gln
		195					200					205			
Gln	Leu	Val	Glu	Ala	Phe	Glu	Glu	Met	Thr	Arg	Asn	Leu	Phe	Ser	Leu
	210					215					220				
Pro	Ile	Asp	Val	Pro	Phe	Ser	Gly	Leu	Tyr	Arg	Gly	Val	Lys	Ala	Arg
225					230					235					240
Asn	Leu	Ile	His	Ala	Arg	Ile	Glu	Glu	Asn	Ile	Arg	Ala	Lys	Ile	Arg
				245					250					255	
Arg	Leu	Gln	Ala	Thr	Glu	Pro	Asp	Gly	Gly	Cys	Lys	Asp	Ala	Leu	Gln
			260					265					270		
Leu	Leu	Ile	Glu	His	Ser	Trp	Glu	Arg	Gly	Glu	Arg	Leu	Asp	Met	Gln
		275					280					285			
Ala	Leu	Lys	Gln	Ser	Ser	Thr	Glu	Leu	Leu	Phe	Gly	Gly	His	Glu	Thr
	290					295					300				
Thr	Ala	Ser	Ala	Ala	Thr	Ser	Leu	Ile	Thr	Tyr	Leu	Gly	Leu	Tyr	Pro
305					310					315					320
His	Val	Leu	Gln	Lys	Val	Arg	Glu	Glu	Ile	Lys	Ser	Lys	Gly	Leu	Leu
				325					330					335	
Cys	Lys	Ser	Asn	Gln	Asp	Asn	Lys	Leu	Asp	Met	Glu	Thr	Leu	Glu	Gln
			340					345					350		
Leu	Lys	Tyr	Ile	Gly	Cys	Val	Ile	Lys	Glu	Thr	Leu	Arg	Leu	Asn	Pro
		355					360					365			
Pro	Val	Pro	Gly	Gly	Phe	Arg	Val	Ala	Leu	Lys	Thr	Phe	Glu	Leu	Asn
	370					375					380				
Gly	Tyr	Gln	Ile	Pro	Lys	Gly	Trp	Asn	Val	Ile	Tyr	Ser	Ile	Cys	Asp
385					390					395					400
Thr	His	Asp	Val	Ala	Asp	Ile	Phe	Thr	Asn	Lys	Glu	Glu	Phe	Asn	Pro
				405					410					415	
Asp	Arg	Phe	Ile	Val	Pro	His	Pro	Glu	Asp	Ala	Ser	Arg	Phe	Ser	Phe
			420					425					430		
Ile	Pro	Phe	Gly	Gly	Gly	Leu	Arg	Ser	Cys	Val	Gly	Lys	Glu	Phe	Ala
		435					440					445			
Lys	Ile	Leu	Leu	Lys	Ile	Phe	Thr	Val	Glu	Leu	Ala	Arg	His	Cys	Asp
	450					455					460				
Trp	Gln	Leu	Leu	Asn	Gly	Pro	Pro	Thr	Met	Lys	Thr	Ser	Pro	Thr	Val
465					470					475					480

005260 28489960 0968482 092500



Tyr Pro Val Asp Asn Leu Pro Ala Arg Phe Thr Tyr Phe Gln Gly Asp  
485 490 495

Ile

(2) INFORMATION FOR SEQ ID NO:33

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: base pairs
- (B) TYPE: nucleic acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33

CGCACCCAG GAGGCGCGCT CGGAGGGAAG CCGCCACCGC CGCCGCCTCT GCCTCGGCGC	60
GGAACAAACG GTTAAAGATT TTGGGCCASC GCCTCCGCGG GGGGAGGAGC CAGGGGCCCC	120
AATCCCGCAA TTAAAGATGA ACTTTGGGTG AACTAATTGT CTGACCAAGG TAACGTGGGC	180
AGCAACCTGG GCCGCCTATA AAGCGGCAGC GCCGTGGGGT TTGAAGCGCT GGCGGCGGCG	240
GCAGGTGGCG CGGGAGGTCG CGGCGCGCCA TGG	273

(2) INFORMATION FOR SEQ ID NO:34

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: base pairs
- (B) TYPE: nucleic acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34

CGCACCCCCA GGAGGCGCGC TCAGAGGGAA GCCGCCAGTG CGCCGCCTCT GCCTCGGCGC	60
GGAACAAACG GTTAAAGATT TTTTGGGCA GCGCCTCGAG GGGGAGGAG CCAGGGGCCC	120
GATCCGCAAT TAAAGATGAA CTTTGGGTGA ACTAATTGT CTGACCAAGG TAACGTGGGC	180
AGTAACCTGG GCGGCCTTAT AAAGAGGGCG CGCGGCGGGG TTCGAGCTA GGGAGGCGGC	240
GGCAGGTGGC GCGGGAGGCT GAAGCGTGCC ATGG	274

(2) INFORMATION FOR SEQ ID NO:35

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 319 base pairs
- (B) TYPE: nucleic acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35

TCGGGGGAAT TAACACCTTT TCAAAGTGAA ATCTCAGGAT TGTCTGCCTT CTACAGGAGG	60
TGGTATTAAA ATGCGCCTAT AACAAATGGT TGAGAGTTTG GAGCCGCTTC TGCCCTGTGG	120
GCGGGGCGAG ATGACACCAC AATTAAAGAT GAACTTTGGG TGAACATAATT TATCTGAGGA	180
AGTTAACAGG AGGAGACCTG CGCGCAATGG ATATATAAGG GCGCGCAGGC GAGGACGCCC	240
TCAGTTTGTG CGTAAAGACG CGTCTCCTCT CCAGAAGCTT GTTTTTCGTT TTGGCGATCA	300
GTTGCGCGCT TCAACATGG	319

005260"28489500



(2) INFORMATION FOR SEQ ID NO:36

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: base pairs
- (B) TYPE: nucleic acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36

GATCCCAGATCTGCCTATTGCGCCCGATGCCCCGAGGCTCTCTTTGGAC  
TCTGGCCCTGAGTTCTTCTGCGCGATCCTTCGGAGACGTCTGGAGGCCTG  
CTTTATGCATCTCTCTTGGACCTCAGTTTCCCCACACGTGGGAGGAGGCA  
GCTGGACGATTCTCTGAAAGGACTTTCCTTGCTTCTCATCACGTGGAAG  
AGAGCCCACCCGGCACCTGGAAATGGAAGCCAGTGAAGGCTGCTTTGGG  
CCGGGGCAKCGGGTGGGACCGGGCGGGAGGGATTCCAAAGAGACCGCCGG  
GAAGGCTAGAGCTTGAATTCGGCTCCTCGGAGTCTTGGCCCTCCCCCA  
CCGCCGCTCGGAGCTCAGCACACCTTGGATGGGGGAGGCGGGCAGCTCC  
TAGCCCCGCACCCAGGAGGCGCGCTCGGAGGGAAGCCGCCACCGCCGCC  
GCCTCTGCCTCGGCGCGGAACAAACGGTTAAAGATTTTGGGCCASCCT  
CCGCGGGGGGAGGAGCCAGGGGCCCAATCCCGCAATTAAAGATGAAGTT  
TGGGTGAAGTAATTGTCTGACCAAGGTAACGTGGGCAGCAACCTGGGCCG  
CCTATAAAGCGGCAGCGCGTGGGGTTTGAAGCGCTGGCGGGCGCGGAG  
GTGGCGCGGGAGGTCGCGGCGCGCATGGGGCTCCCGGCGCTGCTGGCCA  
GTGCGCTCTGCACCTTCGTGCTGCCGCTGCTGCTTCTTGGCTGCGATC  
AAGCTCTGGGACCTGTACTGCTGAGCGGCGCGACCGCAGTTGTGCCCT  
CCCATTGCCCCCGGGACTATSGGSTTCCCCTTCTTTGGGGAAACCTTGC  
AGATGNTACTNCAGGTAAGGGAGGGTGGGGCGGGACAGGCTGCTTCCCCG  
GAGCCCCGCGCGGCTCTGGGCTTCTGCTGAAGTCGGGGTAGGCGCCCCCG  
GGAGGCTAGCTATTGCGGCTAGGAGCAGGGCTGGCGGGAGCGCGCGCTC  
CCCGGMKYMCSCTCAWGCSCRCWWKTMWCCTCCGCTYMCTCCCAMAGCG  
GARSARWKYKGMRGATGAAGCGCAGGAAATACGGCTTCATCTACAAGA  
CGCATCTGTTCCGGCGGCCCCACCGTACGGGTGATGGGCGCGGACAATGTG  
CGGCGCATCTTGCTCGGAGAGCACCAGGCTGGTGTCCGTCCACTGGCCAGC  
GTCGGTGCGCACCATTCTGGGATCTGGCTGCCTCTCTAACCTGCACGACT  
CCTCGCACAGCAGCGCAAGAGGTGGGGGAGGAGCGACGGCTGGACA  
GGGAGGGGGACCCATTTATGAGCGGAATTCGGCTGATGGATGCTAGGC  
GCGGGCTAGCAGCTTGAGGTGGGCTAGGACCTCTGCCAGCTCCAGGTTA  
GCTTTCCAGCTCGGAGAGTGCCATGTGTCTGGCAGGACTGGGGGTGTCT  
GGAAGGGGACGGCGGTAGACGAGAGGGGCGGATGGAGGCTTTTAACGCTG  
TCCCCTCCTCGGGACTCAGGTGATTATGCGGGCCTTCAGCCGCGAGGCAC  
TCGAATGCTACGTGCCGTGATCACCGAGGAAGTGGGCAGCAGCCTGGAG  
CAGTGGCTGAGCTGCGGCGAGCGCGGCTCCTGGTCTACCCCGAGGTGAA  
GCGCCTCATGTTCCGAATCGCCATGCGCATCCTACTGGGCTGCGAACCCC  
AAGTGGCGGGCGACGGGACTCCGAGCAGCAGCTTGTGGAGGCCTTCGAG  
GAAATGACCCGCAATCTCTTCTCGCTGCCCATCGACGTGCCCTTCAGCGG  
GCTGTACCGGTAAGGGCGGCAACGGGCTGCGGACTAGGGGCGCGGGAC  
CTGGGCGTCTGCTCACCGCGCGCGCTCTCTGCGCTCAGGGCATGAAGGC  
GCGGAACCTCATTCACGCGCGCATCGAGCAGAACATTGCGGCCAAGATCT  
GCGGGCTGCGGGCATCCGAGGCGGGCCAGGGCTGCAAAGACGCGCTGCAG  
CTGTTGATCGAGCACTCGTGGGAGAGGGGAGAGCGGCTGGACATGCAGGT  
GAGTAGCAGCTTCAGACCAGGCACTGCGGAGTTTGGTCCCCTGGCTTTCC  
AAGGCGCTGTTCCTGGGGCCCCAAAGCGCGCGCTGGGGCCAGCTTTC  
TGGAGTGGGCGGCGGCTCAGACTACAGCTATGGAATCCCGAAGGAAGGC  
TGAGACACCCGTCAGGAGAGCTGCGGAAGGGGCTGCGGMGGAACCTGGG  
AGCATCCCCTAGCCTTTAMCAGGTTTCAAAGGGAAAGTTGGAATTTGCAA  
AAATGTTAATAAAGAACCTTGCGATTTTAATAAACTAAGACTTTAACTC  
AGGAGTTTCCGGTAGRGGGGTCTGACTCGCCTTACTGCTCCAGCTGAA  
CTAAAGGGACGTTGCATTTTGTAAAGATATTGCTTTCCTTGACTTTCT  
GTCAGCAAAACATTTAGCCCTTCTAGTCTTCCCTCCAGAACTCTCAGTTC  
GATTCTGAGTAATCTTGTCAAACCGCAGGCAGACTTGTGAGAATGTG  
GGTCTCACTCTATTCTTAGGCACTAAAGCAATCTTCAACCGAATCCTCT  
TTGGAGGACACGAAACCAGGCCAGTGCAGCCACATCTCTGATCACTTAC  
CTGGGGCTCTACCCACATGTTCTCCAG

(2) INFORMATION FOR SEQ ID NO:37



(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: base pairs
- (B) TYPE: nucleic acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37

GATCCAGGTTGCTGAAACATATCTCCATATAGGGCAGAACAATTATCAAA  
AGCATAAGAATTGCAGCCACAGCATAGGGAAGAAAGAGGAGTTTTTAAAC  
CACAAACAAAAGGGAGAAAGAGAATTTTAACTTACATTTAATTCAAAA  
GTCTTCAGAGCAACCCGAAACCCTCCTGGAAGTGGGGGATTTCAGTCGAAG  
GGTCTCCTTAATAACACACCCGATGTATYTAAGTTGTTCCAAAATTTCCA  
TGTCCAACCTTGTGTCTTGATTGCTCTTGCAAAGTAAACCCTAYCAAAAY  
AGTCATACAGAGGTGAACAGTYATTTTGTGCTCCAATTAATAATCAGCCCA  
GCAGACGTAAACAGGGCTTAAGTGGAGACTAAACCCAAAGGGCCCATGA  
TGGGAGAGACTGGGAGGGGGGAAACAGCAGCTAATGGCCATTTGCCTGCCC  
AAATCCACTATCTATTTACAATCCCAGGAGAATGCTGCTCACCAGTTAGA  
AGGACCAAGTTTCTCCCCACGCCCCCCCCACCCACACTCACCACCACCAC  
CCACACTAATCAGCTATTACACTATGTATGCCCTTGGACACACCAATTC  
AAGAAAAGTGGAACCTATCTGAGAATCTCCACGGTTCACAAAAGGTGGA  
GGAGGGGTAGGAATACAAGGTCAAACCCTGCCC

(2) INFORMATION FOR SEQ ID NO:38

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: base pairs
- (B) TYPE: nucleic acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38

TCGCGAGGAGCGACCACGGCTTGAAGAGGGGTAGACGAGACCAGATGCTC  
CCCGGCGCCCCCTCATGCGGGTTGCGGTCTCTCTCCTCCACCTCCCTCTC  
AGCGGAGGAAGTTTCTGCAGATGAAGCGCAGGAAATACGGCTTCATCTAC  
AAGACGCATCTGTTTGGGCGGCCACGGTGCGGGTGATGGGCGCGGATAA  
TGTGCGGCGCATCTTGCTGGGAGAGCACCGGTTGGTGTGCGGTGCACTGGC  
CCGCGTGGTGCGCACCATCCTGGGCGCTGGCTGCCCTCTCCAACCTGCAC  
GATTCCTCGCACAAAGCAGCGAAAGAAGGTGAGGGTGAGCTGGCAACTCCT  
TGGCTGGCAGGGAGACCTCATCCTATGGCTTGGTTTCAGGCAAAATAGAAT  
GCGGGGCGAGGGCTAGTCTATGTGGTGGGGACCAGGACCCTCTCTATCT  
GAGATCCACTTTAGCTTTTCTGCTAGCACGTGGGTAGTCTGGGGGGGA  
CTGAAATTCCTGAAAGGGTACTCGGAAAGGCGAAGGGGGGGGGCTGAGG  
GAAAGTAGAGGATTGTAACTACTCTGCTCCTGGGGGGTGCTCAGGTGAT  
TATGCAGGCCTTCAGCCGCGAGGCACTCCAGTGCTACGTGCCCGTGATCG  
CTGAGGAAGTCAGCAGTTGTCTGGAGCAGTGGCTAAGCTGCGGCGAGCGC  
GGCCTCCTGGTCTACCCCGAGGTGAAGCGCCTCATGTTCCGCATCGCCAT  
GCGCATCCTGCTGGGCTGCGAGCCGGGTCCAGCGGGCGGCGGGGAGGACG  
AGCAGCAGCTCGTGAGGCTTTTCGAGGAGATGACCCGCAATCTCTCTCT  
CTTCCCATTTGACGTGCCCTTTAGCGGCCTGTACCGGGTAAGGGCGGTTTG  
CGGAGTCGGAGTAGGGGAACGCAAGCTCGGGCATCCGCTCACCGCCACGC  
TCTCTCCGCGCTCAGGGCGTGAAGGCGCGGAACCTTATACACGCGCGCAT  
CGAGGAGAACATTCGCGCCAAGATCCGCGGGCTTCAGGCTACAGAGCCGG  
ATGGGGGTTGCAAGGACGCGCTGCAGCTCCTGATTGAGCACTCGTGGGAG  
AGGGGAGAGAGGCTGGATATGCAGGTGAGAAGCAATTTCAAAAGGTGCCA  
AGGGCCGGGGAGTGCCCTCTGACTTTCCAGACACACTTTCTGGGGTCTCCA  
AAGCCCTGTCAAGGCCCCAGCTACTTCCAAGTGGGCGGCGATGCTAGGTC  
TAGAGCTTTTCAACCTGTGGGTCTGACCCCTTCACGGAGCCAAACAACC  
CTTTCAGAAGGGTCGCCCTAAGAGCATCTGCATATCCGATATTTACATCAA  
GAAACATAACAGTAGCAAAAATTACCGTTATGAAGTAGCAACAAAGATAAT  
TTTATCGTTGGGGGTCAACCAACACGAGGAACCGTATTAAAGGGTGGCA  
TTGGTCTAGAGAGCTGTGGAAGGGGGTGGCTGAGCAATGGGGAAGATCCC  
AAAGTTCAAAGGGCAAGGCTCATCTACAAAGGTTAAAGCGGAAGAGCAGG  
ATTAAGGGAGTTTTCGCTTTTGTGTTGTTGTTGACTTTCTATGAACA  
AAACGGATTTTACCCTTGAAGTCTTCCGTGCAATATTCTCAGGTCAGGTC  
TTTGTAACAGTGCTATAAACTGCACTCAGATCTGTATAAACTTCCGTTTT



GAAGTGCTTCAGGA